

Fine-Structure Mutational Analysis of a Stage- and Tissue-Specific Promoter Element of the *Drosophila* Glue Gene *Sgs-3*

TAKESHI TODO,[†] MARGARET ROARK,* K. VIJAY RAGHAVAN,[‡] CAROL MAYEDA,
AND ELLIOT MEYEROWITZ

Division of Biology, California Institute of Technology, Pasadena, California 91125

Received 29 May 1990/Accepted 15 August 1990

The *Sgs-3* gene of *Drosophila melanogaster* exhibits a tightly regulated pattern of expression governed by two functionally equivalent elements within 1 kb of the gene, each of which is sufficient to confer third-instar salivary gland-specific transcription. In this report we describe a detailed functional analysis of one of these, the proximal element. To determine the nucleotides responsible for specific expression, we have introduced mutations into the proximal element and then assessed the effects of each alteration on expression in the developing animal. We have identified six particularly important base pairs which are located in two regions separated by nonessential sequences. These base pairs, along with some surrounding sequence, are conserved within the upstream regions of the three glue genes at 68C. Nearly identical groups of base pairs can be found upstream of the other glue genes which have been cloned. This analysis has allowed us to derive a consensus sequence, which we believe contains binding sites for two different factors which interact to direct third-instar salivary gland-specific expression.

The *Sgs-3* gene of *Drosophila melanogaster* encodes a glycoprotein component of the secretion produced in the salivary glands of the third-instar larva, which is expelled just prior to pupariation (1, 18). This "glue" serves to affix the pupa to a surface for the duration of metamorphosis (8). The *Sgs-3* gene is one of three glue genes clustered at 68C on the left arm of the third chromosome (9, 32). Two other of the eight known glue genes, *Sgs-4* at 3C and *Sgs-5* at 90BC, have also been cloned (34). The *Sgs-3* gene exhibits a tightly regulated pattern of expression. It is abundantly transcribed only in a single cell type, of a single tissue, during a single 30-h period of development: the secretory cells of the salivary gland in the mid- to late third larval instar. Our interest has been in understanding how this extraordinarily specific expression pattern is achieved.

Two *trans*-acting regulators of this gene have been previously described (see reference 33 for a review). The steroid hormone ecdysterone must be present at the time of the second molt, or no *Sgs-3* mRNA is produced (14). In addition, particular mutations at the 2B5 locus, such as that of the nonpupariating mutant *npr-1*³, fail to produce *Sgs-3* mRNA (5). Whether either of these factors acts directly at the *Sgs-3* locus is not yet known.

The required *cis*-acting sequences have been investigated in our laboratory and by others (for reviews, see references 29 and 33). An element located within 130 bp of the transcription start site (the proximal element) is sufficient for correctly regulated but low-level expression of this gene (54). A second element, located between -130 and -629 bp relative to the transcription start site (the distal element), is also sufficient for regulated low-level expression (43). When these elements are both present, a 20-fold enhancement of the level of expression is seen. Within the proximal element,

only sequences between -56 and -98 are required for full function (28, 43).

Which base pairs within this 42-bp region are specifically required for the stage- and tissue-specific expression of this gene? Functional analyses of promoters containing point mutations, either naturally occurring or produced *in vitro*, have allowed the identification of the consensus elements required for the regulation of many genes (see, e.g., references 12, 23, 36, and 55). These studies, however, have assessed genes as they are expressed in cultured cells. For a gene which is temporally regulated during development, assessment of the function of regulatory elements in the developing animal is particularly important. Such detailed knowledge of a regulatory element is a prerequisite for eventual isolation of the regulatory protein(s) which binds it (see, e.g., references 22 and 47). Here we report a fine-structure functional analysis of those sequences of the proximal element of the *Sgs-3* gene which have been shown to be sufficient for regulated expression. Individual point mutations and small deletions were introduced into the proximal element, and their effects on expression in the developing animal were assessed. We have identified six particularly important base pairs which fall in two separate required regions. Groups of base pairs within these regions are highly conserved among the glue genes at 68C, and we have found nearly identical groups of base pairs upstream of other cloned glue genes. The functionally important base pairs of the proximal element are conserved in other *Drosophila* species (27) and lie within two sites apparently bound by proteins *in vivo* (P. H. Mathers, Ph.D. thesis, California Institute of Technology, Pasadena, 1989).

MATERIALS AND METHODS

Plasmid construction. The transient assay plasmid pUC119'GOA.98 was constructed as follows. The 3.8-kb *Xho*I fragment, which contains the region from 983 bp 5' of the mRNA start site through the first 12 bp of transcribed sequences of the *Sgs-3* gene and also contains 2.8 kb of transcribed and 3' sequences of the *Adh* gene, was excised from pGOA.98 (43) and subcloned into the *Sal*I site of

* Corresponding author.

[†] Present address: Department of Fundamental Radiology, Faculty of Medicine, Osaka University, Kita-ku, Osaka 530, Japan.

[‡] Present address: Molecular Biology Unit, Tata Institute of Fundamental Research, Bombay 400 005, India.

pUC119' (pUC119' lacks the *EcoRI-SacI-KpnI-SmaI* cloning sites of pUC119). Proximal regions containing mutations were then subcloned into this plasmid as *SalI-KpnI* fragments replacing the wild-type (WT) sequence. The 3.8-kb *XhoI* fragment was subsequently subcloned into Carnegie 20 for transformation into the germ line.

Mutagenesis. Our saturation mutagenesis consisted of two steps. First, point mutations were introduced at random by chemical treatment. Then, the remaining sites at which mutations were not recovered were mutagenized by site-directed mutagenesis.

(i) **Chemical mutagenesis.** For chemical mutagenesis, a 150-bp *SalI-KpnI* fragment containing the sequences between -130 and +12 bp of the *Sgs-3* gene was excised and subcloned into M13 mp18 or mp19. Single-stranded DNA (ssDNA) derived from this phage was then used as a substrate for chemical mutagenesis. To maximize the chance of obtaining single-base substitutions in all positions of the target DNA, three different chemicals were used: nitrous acid (deamination of deoxycytosine, deoxyadenine, and deoxyguanine), formic acid (depurination), and hydrazine (breakage of pyrimidine rings) (35).

(a) **Chemical reactions.** Chemical reactions were performed essentially by the method of Meyers et al. (35). About 40 μ g of ssDNA was suspended in 40 μ l of TE (25) and either 10 μ l of 2.5 M sodium acetate, pH 4.2, and 50 μ l of 2 M sodium nitrite; or 60 μ l of concentrated formic acid (18 M); or 60 μ l of concentrated hydrazine (12 M) was added, each reaction having a final volume of 100 μ l. After incubation at room temperature (RT) for 69, 12, and 12 min, respectively, 200 μ l of 2.5 M sodium acetate (pH 5.5), 100 μ l of H₂O, and 1 ml of ethanol were added, and the reactions were centrifuged to pellet the DNA. After two more ethanol precipitations, the pellet was suspended in 80 μ l of TE.

(b) **Reverse transcription of chemically treated templates.** To each chemically treated ssDNA sample (80 μ l), 10 μ l of 10 \times KS (70 mM Tris hydrochloride [pH 7.5], 70 mM MgCl₂, 0.5 M NaCl, 20 mM dithiothreitol) and 9 μ l of 20- μ g/ml oligonucleotide primer (M13 universal primer: 5'-GTAAAA CGACGGCCAGT-3') were added; the reactions were heated to 75°C for 5 min, and then annealed at 40°C for 15 min. Ten microliters of deoxynucleoside triphosphate mix (2 mM in each deoxynucleoside triphosphate) and 40 U of avian myeloblastosis virus reverse transcriptase were added, and the reaction was incubated at 40°C for 1 h. After extraction with phenol-CHCl₃ and precipitation with ethanol, the *SalI-KpnI* 150-bp fragment was excised and substituted for the *SalI-KpnI* fragment of the pUC119'GOA.98 assay vector (Fig. 1).

(c) **Screening of point mutants.** After transformation and plating, each colony was amplified in a microtiter well, and 5 μ l of culture was spotted onto a nitrocellulose filter. These filters were treated with 10% sodium dodecyl sulfate for 3 min, denatured with 0.3 M NaOH-1.5 M NaCl for 5 min, neutralized with 1 M Tris hydrochloride (pH 7.4)-2.5 M NaCl for 5 min, and baked for 2 h at 80°C in a vacuum oven. Nine replica filters were made in this manner, and each filter was probed with one of nine different ³²P-labeled oligonucleotide (20-mer) probes, each of which spans part of one strand of the 100-bp region. The nine probes together cover both strands of the 100-bp region (-130 to -30 bp) where mutations were desired (Fig. 1). The filters were prehybridized in 6 \times SSPE (25)-10 \times Denhardt solution (1 \times Denhardt solution is albumin-polyvinylpyrrolidone-Ficoll, each at 0.2 mg/ml)-0.2% sodium dodecyl sulfate-1 mM EDTA (pH 8.0) for 1 h at 42°C and hybridized overnight at RT in the same

solution plus end-labeled oligonucleotide probe (usually 300 to 600 cpm/pg and 10⁷ cpm/ml). In the washing procedure, to control the stringency of the hybridization, a solution of tetramethylammonium chloride (Me₄NCl) was used. Me₄NCl binds selectively to A-T base pairs, thus displacing the dissociation equilibrium and raising the melting temperature. At 3 M Me₄NCl, this displacement is sufficient to shift the melting temperature of A-T base pairs to that of G-C base pairs. When filters hybridized with oligonucleotide probes are washed in this solution, the melting temperature is independent of base composition and dependent only on probe length (56). The filters were rinsed three times in 6 \times SSPE and twice for 30 min in 6 \times SSPE at RT. The filters were then rinsed with the Me₄NCl wash solution (3.0 M Me₄NCl-50 mM Tris hydrochloride [pH 8.0]-2 mM EDTA containing sodium dodecyl sulfate at 1 mg/ml) at 37°C and washed twice for 20 min in Me₄NCl wash solution at 59°C. At this temperature, clones having as little as a single-base alteration will result in the probe melting off. Absence of hybridization of a particular colony to a particular probe identifies the region of the mutations. Since the probes cover both strands (Fig. 1), each 10-bp region is covered by two probes (an upper- and a lower-strand probe). Colonies which failed to hybridize to two overlapping probes were picked from the master dish; single-stranded plasmid DNA was prepared by using the method of Vieira and Messing (53), and the ssDNA was sequenced by the dideoxy chain termination method (45) to verify the change introduced. A total of 2,150 colonies were screened, and 96 mutants were identified (77 from nitrous acid treatment, 10 from hydrazine, and 9 from formic acid). As some proved to have the same base change, mutations in a total of 46 of the 75 sites between position -106 and -32 bp of *Sgs-3* were isolated by using this method. Of these 46 mutants, 20 have two base changes (double mutants) and 5 have three base changes (triple mutants).

(ii) **Site-directed mutagenesis.** After chemical mutagenesis, unmutated sites remain at random. To target these sites, a series of 20-base oligonucleotides were made. Each oligonucleotide covered four or five unmutated sites. At the time the oligonucleotides were synthesized, mutant and WT residues were mixed at each site, so that each oligonucleotide was a mixture of WT sequence and mutant sequence at these sites. These mixed oligonucleotides were used as primers for site-directed mutagenesis, by using the method of Kunkel (20).

(a) **Uridine-containing template.** As the template for mutagenic oligonucleotide-primed DNA synthesis, ssDNA derived from transient assay plasmid pUC119'GOA.98 or M13 phage mp18 into which the 150-bp *SalI-KpnI* fragment was subcloned was used. To prepare uracil-containing ssDNA from the pUC119-derived plasmid, transient assay plasmid pUC119'GOA.98 was used to transform RZ1032 (HfrKL16, dut1, ung1; [21]). Single colonies were picked and cultured in YT medium (yeast extract, 5 g/liter; Tryptone, 8 g/liter; NaCl, 5 g/liter) containing 0.25 g of uridine per ml. Single-strand phage production was initiated by superinfection of a mid-log phase culture with M13KO7 helper phage at a multiplicity of infection of 2 to 10. (53). After 1 h of incubation at 37°C, kanamycin was added to a final concentration of 70 g/ml. After a further 16 h of incubation, supernatant was harvested and ssDNA was isolated. M13 mp18 phage with uracil-containing DNA were prepared as previously described (20).

(b) **Mutant strand synthesis.** Uracil-containing ssDNA (0.4 pmol) was annealed to 4 pmol of kinased mutagenic oligo-

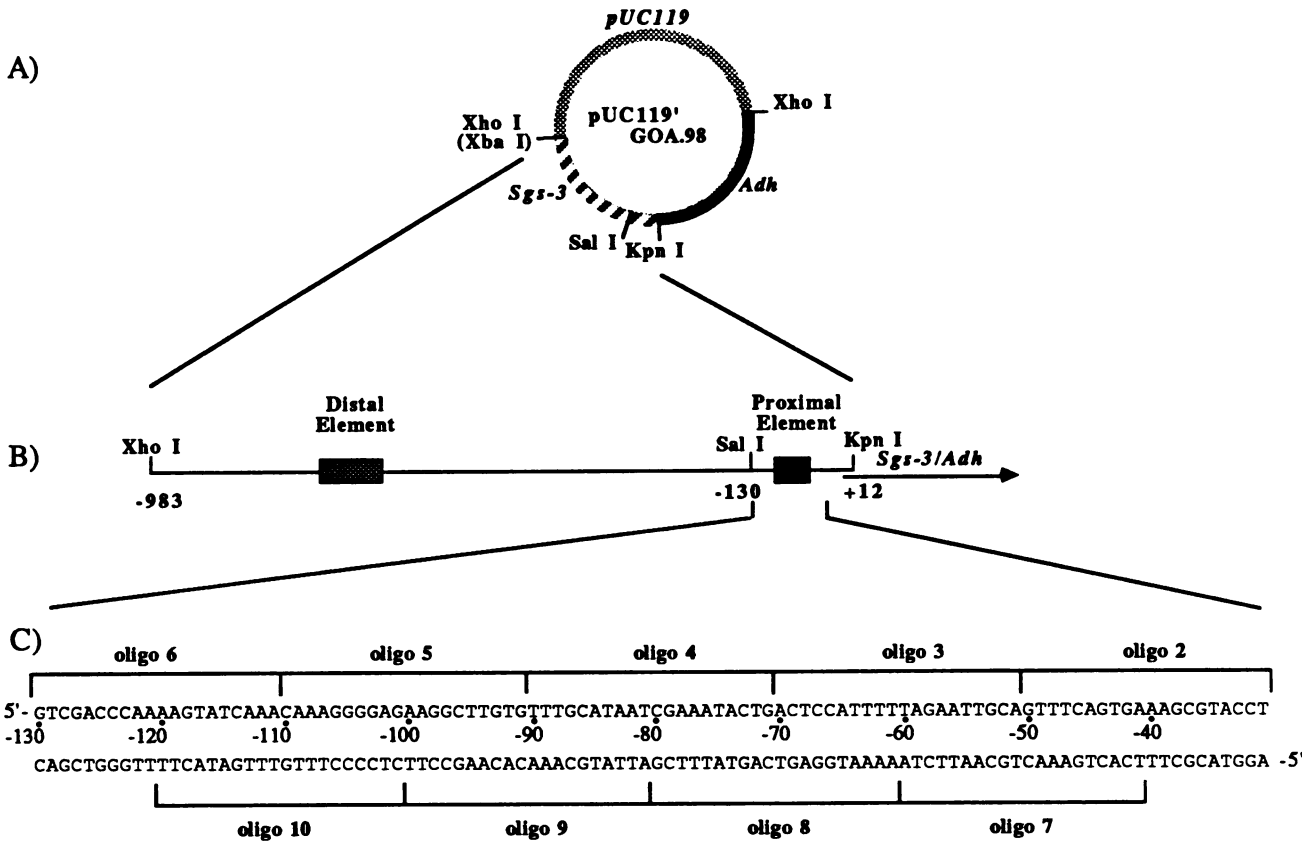


FIG. 1. (A) The upstream *Sgs-3-Adh* hybrid gene construct pUC119'GOA.98 (see also Materials and Methods). The WT sequence has an *Xba*I site at -983, where this construct has an *Xho*I site. Each mutation to be tested was subcloned into this plasmid as a *Sal*I-*Kpn*I fragment. The 3.8-kb *Xho*I fragment containing the fusion gene was subsequently subcloned into Carnegie 20 for testing in the germ line. (B) Schematic map of the fusion gene. The positions of the proximal and distal regulatory elements of *Sgs-3* referred to in the text are indicated with hatched boxes. Numbers indicate the positions of the restriction sites shown, with reference to the start site of transcription of *Sgs-3*. (C) Nucleotide sequence of the region, surrounding the proximal element of *Sgs-3*, which was screened for mutations. Numbering is the distance in base pairs upstream, where -1 is the first base pair upstream of the start site of transcription. Oligos 2 through 10 were used as probes in the screen for point mutations. Clones which failed to hybridize to a probe were assumed to have a mutation in the region covered by that probe and so were picked and sequenced. See Table 2 for a list of the mutations identified.

nucleotide by incubation for 3 min at 70°C followed by incubation for 15 min at RT in the presence of 10 mM Tris hydrochloride (pH 7.5) and 10 mM MgCl₂. Second-strand synthesis and ligation were catalyzed by 5 U each of T4 polymerase and T4 ligase in the presence of 20 mM HEPES (*N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid), (pH 7.8); 2 mM dithiothreitol; 10 mM MgCl₂; 500 M each dATP, dTTP, dGTP, and dCTP; and 2 mM ATP. This mixture was incubated first for 5 min on ice and then for 5 min at RT and 1 h at 37°C (21), and 10% of this reaction mixture was used to transform JM101. As the primer used was a mixture of WT and mutant sequences, the resulting plasmids and phage were a mixture of WT and mutants. Colonies or plaques were then screened to identify mutant clones by using the same method described above for chemical mutagenesis.

(iii) **Deletion mutants.** Internal deletion mutants were made by using the same site-directed mutagenesis protocol described above, with 28-base oligonucleotides containing the deletion as primer.

All deletion mutants are listed in Table 1; all point mutants identified and tested for expression of alcohol dehydrogenase (ADH) in the transient assay are listed in Table 2.

Transient expression assay. This method was originally developed by Martin et al. (30). The details of the procedure

as used in this work have been previously described (43). Embryos were injected with plasmid pUC119'GOA.98, containing the mutant proximal element to be tested, at 0.5 mg/ml in injection buffer (49). Third-instar larvae which developed from the injected embryos were dissected in cold Tris-buffered Ringer solution (TBR) (10 mM Tris [pH 8.0], 80 mM NaCl, 5 mM KCl, 1.4 mM CaCl₂), fixed for 5 min in cold Tris-buffered Ringer solution containing 1% glutaraldehyde,

TABLE 1. Number of glands stained in transient assays of fusion genes containing small deletions

Deletion endpoints ^a	Expression ^b
-106, -98	12/17
-101, -93	5/17
-92, -84	0/14
-83, -75	2/6
-74, -66	0/21
-62, -54	6/7
-60, -52	5/10

^a For each construct, the base pairs listed are deleted as well as the intervening base pairs.
^b All staining was with *sec*-butanol. Ratios shown are the number of salivary gland pairs showing any cells stained to the total number of pairs assayed. Ratios are not quantitatively significant (see Materials and Methods).

TABLE 2. Numbers of glands stained in transient assays of fusion genes containing point mutations in the proximal element

Position	WT base	Mutant base	Expression ^a	
			Isobutanol	sec-butanol
-103, -53	G, G	A, A	14/20	X
-102	A	T	X	10/12
-101, -75	G, T	A, C	5/28	X
-100, -32	A, C	G, G	X	3/6
-99, -51	A, A	G, G	1/7	X
-98	G	C	X	8/8
-97	G	T	2/5	X
-96	C	T	1/5	X
-95	T	A	X	11/18
-94	T	C	13/16	X
-93	G	A	4/26	X
-92, +2	T, T	C, C	1/20 ^b	X
-91, -103	G, G	A, A	4/10	X
-90	T	A	X	0/16
-89	T	C	0/11	0/18
-89, -73	T, C	C, T	0/21	X
-88	T	C	0/27	0/16
-87	G	C	X	0/24
-86	C	T	3/20	X
-85	A	G	2/4	X
-84, -43	T, T	C, C	4/5	X
-83	A	T	X	4/10
-82	A	G	3/15	X
-81	T	C	2/5	X
-80	C	T	1/10 ^c	X
-79	G	C	X	2/13
-78	A	G	2/3	X
-77	A	T	2/10	X
-76	A	T	4/8	X
-75	T	C	18/29	X
-74, -86	A, C	G, T	3/47 ^c	X
-73	C	G	X	5/7
-72	T	A	4/16	X
-71	G	T	3/35 ^c	X
-70	A	T	X	1/8
-69	C	T	1/7	X
-68	T	A	0/12	3/28 ^d
-67	C	T	1/19 ^b	7/31 ^d
-66, +1	C, A	T, G	5/40	X
-65	A	G	6/19	X
-64	T	A	X	1/4
-63	T	G	13/16	X
-62	T	C	2/8	X
-61, -55	T, T	C, G	X	5/26
-60	T	A	X	3/15
-59	A	T	X	3/5
-58, -81	G, T	A, C	X	2/6
-54	T	A	X	1/2
-52	C	G	X	3/3
-47	T	C	X	4/8
-45, -67	A, C	G, T	0/8	3/9 ^d
-44, -43, -36	G, T, G	A, C, A	6/6	X
-43	T	C	28/49	X
-40	A	T	X	11/14
-37	C	G	X	1/9 ^b
-35, +2, -93	T, T, G	C, C, A	3/17	X
-34	A	G	5/20	X

^a Ratios shown are the number of salivary gland pairs showing any cells stained to the total number of pairs assayed.

^b Showed a low ratio of stained glands together with a low level of staining in the few cells stained and was therefore classified as \pm .

^c Gave a low ratio of stained glands but are classified as + because cells are stained strongly in those glands that stained.

^d Isobutanol is a poorer substrate for *Drosophila* ADH, so some constructs which are negative with this alcohol are positive with *sec*-butanol.

washed extensively in cold Tris-buffered Ringer solution (usually three changes, overnight), and then stained in 0.2 M Tris (pH 8.0)–0.5 mg of NAD⁺ per ml–0.02 mg of phenazine methosulfate per ml–0.5 mg of Nitro Blue Tetrazolium per ml and 3.2% iso- or *sec*-butanol for 15 to 30 min. Following staining, glands were placed in 70% ethanol–30% acetic acid to stop the staining reaction, rinsed in 100% ethanol, and mounted in Permount.

The results of transient expression assays of point and deletion mutants are summarized in Tables 2 and 3 and Fig. 2. Previous studies using this transient assay system (43) had shown that the absolute number of stained glands (i.e., the ratio of stained to total glands) is not quantitatively significant. There are probably a number of reasons for this, most importantly the variable efficiency of the injection procedure (varying amounts of DNA injected, numbers of surviving embryos, etc.). In those experiments, constructs expressing >50% of WT activity showed some darkly stained cells, usually at a ratio of 1/10 or more. Therefore, for the present analysis, if no stained glands were seen, that construct was classified as negative (–). Constructs were scored as positive (+) if any of the glands showed some cells stained a dark purple. As before, the absolute number of glands with dark staining was not considered quantitatively significant. For instance, a point mutation at –80 bp [PM(–80)] gave only one pair of stained glands out of 10 pairs dissected, but this pair was darkly stained, so this mutation is classified as positive (Fig. 3A). If only occasional cells (that is, a low ratio) were seen to stain and that staining was a very light color or if a construct was negative with isobutanol but positive with *sec*-butanol, then that construct was classified as a low expresser (\pm). For example, with isobutanol, PM[–67] gave one lobe with two weakly stained cells out of 19 pairs of glands dissected; this was classified as low (\pm) (Fig. 3B). Low expressers never show dark staining with isobutanol, and negatives never show staining with either alcohol in this assay.

Germ line transformation. Germ line transformation followed the standard methods (44, 49) and was described previously (43).

Spectrophotometric assay of *Adh* activity. ADH enzyme activity was measured by using the method of Sofer and Ursprung (48), and details were described previously (43). Briefly, crude extracts were prepared from dissected salivary glands, and portions of this extract were added to a reaction mix containing substrate (2-butanol) and cofactor (NAD⁺). Change in A_{340} (due to NADH) was measured over time: a unit of ADH enzyme activity was defined as $\Delta A_{340} = 0.001/\text{min}$. Total protein was also measured (by the method of Lowry [24]), and the data was expressed as units of ADH activity per μg of total protein.

Direct sequencing of transformed genes. Genomic DNA was prepared from adult flies by the method of Bender et al. (2), purified by banding in CsCl₂, and quantified spectrophotometrically. The *Adh-Sgs-3* fusion gene was amplified specifically by a polymerase chain reaction with one primer specific for *Sgs-3* upstream sequences (5' [bp –192]–CGGGTATCTGTTAGTCTCGTTAGCG–3' [bp –167]) and a second primer specific for the *Adh* gene-coding sequences (5' [+302]–CTTTGGATTGATTGCCTTCAGCTCG–3' [+277]). Exponential amplification was performed in a 50 μl reaction mix containing 50 mM KCl, 1 mM MgCl₂, 10 mM Tris (pH 8 to 8.5), 0.1% Triton X-100, 50 μM each deoxynucleoside triphosphate (≈ 40 ng per reaction for a 25-mer), 0.1 μM each primer, 25 ng of genomic template per ml, and 2.5 to 3 U of *Taq* polymerase (Promega) overlaid with 50 μl of mineral oil.

TABLE 3. ADH enzyme activity determined by spectrophotometric assay.

Transformant	Units of ADH enzyme activity per μg of total protein ^a					%
	Measurement 1	Measurement 2	Measurement 3	Mean	$\bar{X} \pm \text{SEM}$	
Tf(2)PM[−90]−1	0.071	0.068	0.069	0.069	0.15 \pm 0.04	4
Tf(2)PM[−90]−2	0.227	0.170	0.189	0.195		
Tf(3)PM[−90]−3	0.139	0.134	0.245	0.173		
Tf(2)PM[−89]−1	0.793	0.697	0.763	0.751	0.68 \pm 0.12	20
Tf(3)PM[−89]−2	0.630	1.194	0.695	0.840		
Tf(2)PM[−89]−3	0.410	0.439	0.460	0.436		
Tf(3)PM[−88]−1	0.800	2.146	0.323	1.090	0.68 \pm 0.23	20
Tf(3)PM[−88]−2	0.936	0.308	0.802	0.682		
Tf(2)PM[−88]−3	0.452	0.122	0.260	0.278		
Tf(3)PM[−87]−1	1.932	2.177	2.243	2.117	1.49 \pm 0.31	40
Tf(2)PM[−87]−2	1.782	0.810	0.946	1.179		
Tf(2)PM[−87]−3	1.279	1.411	0.817	1.169		
Tf(3)dm[−92;−84]−1	1.017	0.672	0.566	0.752	0.81 \pm 0.07	20
Tf(2)dm[−92;−84]−2	1.064	0.433	0.697	0.732		
Tf(3)dm[−92;−84]−3	0.964	0.828	1.064	0.952		
Tf(3)PM[−68]−2	nd ^b	3.701	3.936	3.818	3.27 \pm 0.55	90
Tf(3)PM[−68]−3	3.346	3.111	1.716	2.724		
Tf(2)PM[−67]−1	2.573	3.186	1.955	2.572		
Tf(2)PM[−67]−2	0.671	0.370	0.282	0.441	1.55 \pm 0.62	40
Tf(3)PM[−67]−3	4.160	0.333	0.403	1.632		
Tf(2)dm[−74;−66]−1	0.407	0.948	0.526	0.627		
Tf(3)dm[−74;−66]−2	1.727	2.132	1.758	1.873	1.32 \pm 0.37	40
Tf(2)dm[−74;−66]−3	1.148	1.001	2.222	1.457		
Tf(3)GOAX.98−2	2.738	2.436	5.314	3.50 \pm 0.91		100

^a Measured in the salivary glands of larvae homozygous for transformed fusion genes containing point mutations or small deletions.^b ND, Not done.

Thirty cycles (45 s at 94°C, 45 s at 56°C, 1 min at 72°C) were performed in a TempCycler (model 50, Coy Laboratory Products, Inc.). The aqueous phase was then carefully removed, and a 5- μl portion was used as template in a linear amplification reaction. This second reaction mixture was the same as for exponential amplification, except for the inclusion of only a single primer (in this case, the *Adh* gene-specific primer) to generate ssDNA; the temperature cycles were also the same. Several 50- μl linear reactions were performed and then combined, and the ssDNA was purified by diluting and filtering in a Centricon 30 unit (Amicon), following instructions of the manufacturer. The <50- μl sample was then dried in a Speed Vac concentrator (Savant Instruments, Inc.) and sequenced by the chain termination technique, by using a nested primer specific for *Sgs-3* upstream sequences (5'[bp −146]-AAGTCTAGGCGATC GAGTCG-3'[bp −127]). A Sequenase kit and [³⁵S]dATP were used after suspending the dried ssDNA in H₂O for the annealing reaction, following the instructions of the manufacturer. Autoradiography was overnight at RT with pre-flashed XAR-5 film (Kodak).

RESULTS

The *in vivo* functional assay. Our assay for function of the proximal element makes use of a hybrid gene in which 5' upstream sequences of the *Sgs-3* gene are attached in the untranslated leader sequence to create a transcriptional fusion to a reporter gene coding for *Drosophila* alcohol dehydrogenase (*Adh*). Expression of this fusion gene can be assayed histochemically or spectrophotometrically. Since we wished to test a relatively large number of mutant sequences, we made use of the transient assay of Martin et al. (30) (see also reference 43). In this assay, plasmid DNA is injected directly into preblastoderm embryos, and subsequently, third-instar salivary glands of the injected animals are stained histochemically for ADH activity. Compared with germ line transformation, this assay is very rapid, though not quantitative (see Materials and Methods for a detailed description of this assay and its interpretation). Genes including only the proximal element of *Sgs-3* attached to the *Adh* gene, which express at low levels in the germ line, are negative for ADH activity in the less-sensitive rapid

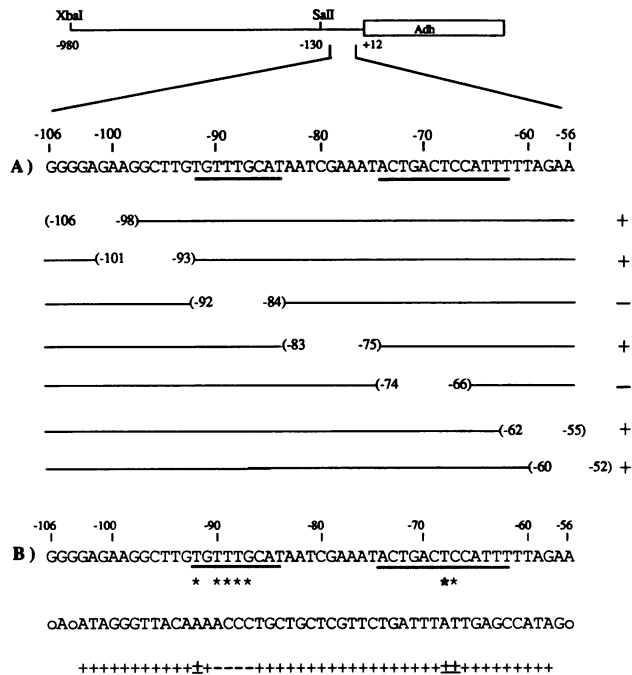


FIG. 2. Expression of 9-bp internal deletion mutants and of point mutants in the transient assay. The diagram at the top shows the relevant portion of pUC119'GOA.98, including the upstream sequences of *Sgs-3* which contain the distal and proximal elements (thin line), and the *Adh* gene (box). (A) The nucleotide sequence of the proximal element, with the locations of deletions listed below. Heavy lines underneath the sequence indicate those nucleotides shown to be important by deletion analysis. The thin lines indicate the *Sgs-3* sequence remaining in the internally deleted plasmids. Parentheses indicate the deleted 9-bp region, and the endpoints of the deletions, which are inclusive, are indicated inside. Plasmids were named from the deletion endpoints. Each plasmid was tested for its ability to direct salivary gland expression of ADH activity by using the transient assay (see Materials and Methods), and the results are shown to the right. A + indicates that histochemical staining was seen in the salivary glands of some injected animals; - indicates that no staining was seen. The results are listed in more detail in Table 1. (B) The WT nucleotide sequence of the proximal element is shown again; asterisks indicate those nucleotides shown to be important in the point mutation analysis (see below). Below the WT sequence, base changes present in individual point mutants are shown (see Table 2 for a detailed listing). A O indicates that a mutation was not recovered at that site. Each point mutant was tested for its ability to direct salivary gland expression of ADH activity by using the transient assay, and the results are shown underneath. A + indicates that histochemical staining was seen in the salivary glands of some injected animals; - indicates that no staining was seen; \pm indicates that low or no staining was seen with isobutanol, but staining was seen if *sec*-butanol was used as substrate. The point mutant at -68 was later shown to express at a high level in the germ line; this is indicated by a boldface asterisk under the WT sequence.

assay. Therefore, we included in our test plasmid the distal regulatory element, which when present together with the proximal element gives a greater than 20-fold increase in expression: that is, it acts as an enhancer (43). The fusion gene with 983 bp of *Sgs-3* 5' sequence present, including both regulatory elements, is positive in the transient assay. In the context of this higher expression level, changes which significantly decrease functioning of the proximal element will lead to a negative result: failure to stain for ADH activity. The test plasmid is named pUC119'GOA.98 and is



FIG. 3. Somatic expression of ADH activity in the transient assay. Glands were dissected from animals injected with *Sgs-3*-*Adh* fusion genes containing point mutations and fixed and stained by using isobutanol as the substrate. See Materials and Methods for details of the assay. (A) Animals were injected with PM(-80); this mutation was classified as +. (B) Animals were injected with PM(-67); this mutation was classified as \pm . Bar = 0.1 mm.

shown in Fig. 1 (see also Materials and Methods). Each mutation to be tested was subcloned into this test plasmid, replacing the WT sequence of the proximal element with the mutant sequence.

Internal deletion mutants of the proximal element. Our previous work has shown that the proximal element lies between -106 and -56 bp 5' of the *Sgs-3* mRNA start site (43). Since the binding sites of regulatory proteins are usually small, in the range of 8 to 20 bp, we predicted that only a portion of this element would contain required sequences. As a step toward narrowing down the required sequences, a series of 9-bp internal deletion mutations were created by site-directed mutagenesis to cover these sequences, and each was tested for expression in the transient assay. When five of the seven deletions were tested, staining for ADH activity was seen, indicating that none of these greatly affected functioning of the proximal element. However, two deletion mutations, from -92 to -84 bp and from -74 to -66 bp [dm(-92/-84) and dm(-74/-66), respectively] showed no detectable expression of ADH activity in the transient assay (Fig. 2A and Table 1). Mutation dm(-65/-57) shows expression; however, as the sequence ATT at -65 to -63 is repeated at -56 to -54, this deletion restores the WT sequence at -65 to -63. Since mutation dm(-60/-52) also shows the expression of ADH, the ATT at -56 to -54 is not responsible for *Sgs-3*-specific expression; however, the base pairs at -65 to -63 remain to be tested by deletion (although they have been tested in point mutations; see below and Fig. 2B, Table 2). These results indicate that two small regions from -92 to -84 and from -74 to -63 contain sequences which are particularly important for function of the proximal promoter element. Since the deletion

dm(-83/-75) does express, the 9 base pairs between these two regions are not required for expression, and the exact spacing between the two required regions is not crucial. These two small regions are marked with a bold line under the WT nucleotide sequence in Fig. 2A.

Point mutations of the proximal element. The deletion analysis narrowed the required sequences to a subset of the base pairs of the proximal element. On the basis of these results, we predicted that point mutations in some of the base pairs falling within these two small regions would negatively affect expression of the fusion gene and conversely, that no point mutations falling outside these regions would have a detectable effect. Therefore, to map the functional parts of the proximal element more precisely, point mutations in 48 of the 51 bp of the element were made and recovered in vitro and cloned into the test plasmid. Of these, 46 have been tested for expression in the transient assay.

Two ADH assays were used to analyze expression of the point mutations, by using two different substrates for the enzymatic reaction. The ADH enzyme of *D. Melanogaster* shows a marked substrate preference for secondary alcohols: the rate of oxidation is up to 100 times higher for secondary alcohols than for primary alcohols (48). This means that histochemical staining reactions performed with *sec*-butanol have a lower threshold than those performed with isobutanol; that is, fusion genes with a lowered level of expression may test positive with *sec*-butanol but be negative with isobutanol. For this reason, *sec*-butanol is too sensitive for screening mutations which have only a slightly reduced level of expression. To identify such moderately defective mutations, some point mutations, mapping between -92 and -84 and between -74 and -63 bp, were screened first using isobutanol.

Each mutation tested which mapped outside the -92 to -84 and -74 to -63 regions, which were shown to be necessary in the deletion analysis, showed expression (see Table 2 and Fig. 3A). Also, within these two regions, many of the base pair changes tested allowed expression. However, seven mutations showed either no expression of ADH or convincingly reduced expression of ADH (Fig. 2B and Table 2). When assayed with the less sensitive substrate isobutanol, PM(-89), PM(-89/-73), PM(-88), and PM(-68) showed no salivary gland expression of ADH, and PM(-92/+2) and PM(-67) showed very weak salivary gland expression of ADH (see Fig. 3B for an example). [The mutation at +2 in PM(-92/+2) is not responsible for the low expression of this gene, because the triple mutation PM(-35/+2/-93) expresses at a high level.] With the more sensitive substrate *sec*-butanol, PM(-90), PM(-89), PM(-88), and PM(-87) still showed no expression of ADH. Together these results indicate that mutations at four sites (-90, -89, -88, and -87) strongly affect the functioning of this promoter, and mutations at three sites (-92, -68, and -67) moderately affect expression. Thus, sequences within only two regions of the proximal element, (-92)T-TTTG(-87) and (-68)TC(-67), are shown to be particularly important for *Sgs-3* expression, and our predictions are confirmed. These seven sites are marked with asterisks under the wild-type nucleotide sequence in Fig. 2B.

Quantitation of expression of mutants in germ line transformants. It is not possible to measure the level of expression of a particular construct quantitatively by using histochemical staining of salivary glands in a transient assay; the signal is expressed in a mosaic fashion, with variable numbers of cells expressing the reporter gene in different experiments. How-

ever, once germ line transformants have been established, the level of expression can be determined more accurately by measuring the ADH enzyme activity in crude cell extracts made from salivary glands of the transformants (see Materials and Methods). Eight of the mutant constructs which showed a lower level of expression than WT in the transient assay were introduced into the germ line of *Adh* mutant flies. For each construct, three (or in one case two) independent transformant lines were analyzed for expression of the fusion gene. Each mutation-bearing gene expressed some ADH activity, as seen by histochemical staining of tissues of transformed animals (data not shown); this was expected, since the distal element alone can drive expression of attached genes. This expression was seen only in the salivary glands of the third-instar larva. The levels of expression of the mutant genes in these transformants was measured by using a spectrophotometric assay for enzyme activity, and the results of this analysis are presented in Table 3. The construct containing a deletion which removes the first region of required base pairs, dm(-92/-84), expresses at about 20% of WT level. WT level, or 100%, for this analysis is defined as the level expressed by GOAX.98, which has the WT *Sgs-3* sequence from +12 to -983, which is also present in pUC119/GOA.98 (43). Since constructs with the entire proximal element deleted express only 5% of WT levels of enzyme activity (43), the proximal element in dm(-92/-84) appears to still retain some function. A similar result is found for constructs with the second region of required base pairs deleted, dm(-74/-66); these express about 40% of WT level. Therefore, we find that neither region of the element is wholly responsible for the function of the proximal element. This is consistent with previous results (28), in which residual function was found in each of the 20-bp deletions of the proximal element tested. The point mutations show a similar picture. PM(-89) and PM(-88) each give about 20% of WT expression, and so are each about as serious a lesion as the deletion which covers them. PM(-87) is less serious, giving approximately 40% of WT. Similarly, PM(-67), at approximately 40% of WT, has about as serious an effect as the deletion which covers it. Surprisingly, PM(-68) expresses more than 90% of WT levels of ADH activity, so it is not clear why this mutation was picked up in the transient assay screen. (The sequence of this mutation was confirmed by direct sequencing; see below [data not shown]). The most curious result, however, is that for PM(-90). This construct expresses only 4% of WT levels of ADH; this is much less than dm(-92/-84), the deletion which removes this region of sequence, and is comparable to levels expressed by a construct with the entire proximal element deleted (TfGOAX.98PA88; see reference 43). We confirmed the identity of this transformed sequence to be sure it contains only the expected mutation. This was done by direct sequencing of the fusion gene using polymerase chain reaction-amplified DNA (see Materials and Methods). Figure 4 shows that the transformed animals carry the expected T-to-A transversion.

DISCUSSION

Sequence homologies at 68C. In the proximal element we have found six positions that are sensitive to single-base substitutions, which fall in the two regions where deletions have an effect. This implies that the regions including these six positions contain *cis*-acting sequences which constitute the regulatory element. For the following reasons, we predict that two other regulatory elements, the distal element of

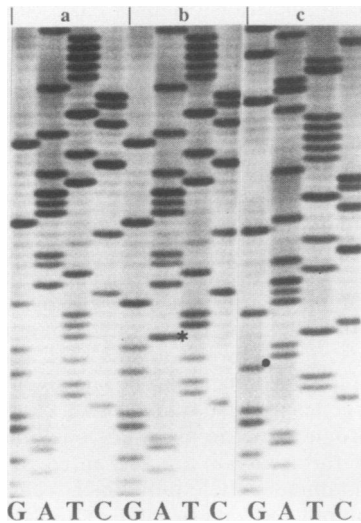


FIG. 4. Sequence of the fusion genes in three transformant lines. The sequence ladders were obtained from polymerase chain reaction-amplified DNA of flies transformed with (a) GOA.98-2, which has a WT sequence (see reference 42), (b) PM(-90)-1 (Fig. 3 and Table 2), or (c) dm(-92;-84) (Fig. 2 and Table 1) (see also Materials and Methods). A* indicates the T-to-A substitution at -90; · indicates the position of the 9-bp deletion between -93 and -83.

Sgs-3 and the regulatory element of the other two glue genes found at 68C, *Sgs-7* and *Sgs-8*, have homologous *cis*-acting sequences.

At least three regulatory elements required for full-level expression of *Sgs-3* have been identified: the proximal element (at -56 to -98 bp), the distal element (at -130 to -629), and one or more elements further upstream (29, 43). We have previously shown that the distal and proximal elements are functionally equivalent; each element alone directs low-level but correct tissue- and stage-specific expression, and when combined, they act together to give a greater than 20-fold increased level of expression; that is, they act synergistically (43). This suggests that both elements may be bound by the same factor(s) and might well have the same or nearly the same sequences for regulation of expression.

Sgs-7 and *Sgs-8* are located upstream of *Sgs-3* in the 68C glue gene cluster and are transcribed as a divergent pair with 475 bp between their 5' ends. A regulatory element required for expression of both these genes resides between -43 and -211 bp 5' of the *Sgs-7* transcription start site (M. D. Garfinkel, Ph.D. thesis, California Institute of Technology,

TABLE 4. Sequences with homology to the proximal element found upstream of other cloned glue genes

Gene or Consensus	Sequence
<i>Sgs-3</i>	(-93)TGTTTG ... (18 bp) ... TCCATT(-62) ^a
	(-624)TATTTG ... (28 bp) ... TCCTTA(-583)
<i>Sgs-7/8</i>	(-108)GGTTTG ... (39 bp) ... TCCATT(-55)
	(-138)TGTTTG ... (31 bp) ... TCCAGT(-95)
<i>Sgs-4</i>	(-324)TATTTG ... (29 bp) ... TCCTAC(-282)
	(-458)ATTTTG ... (31 bp) ... TCGAGT(-502) ^b
	(-545)TATTTG ... (31 bp) ... TCCACA(-501)
<i>Sgs-5</i>	(-119)TGTTTG ... (39 bp) ... TCAATA(-67)
Consensus sequence	... TNTTTG ... (18 to 39 bp) ... TCCAN _A ...

^a Numbers refer to the position of the next base pair before or after the homology, where -1 is the first base pair upstream of the RNA start site for the respective genes.

^b Note that this homology has an orientation, with respect to the transcription start site, which is opposite to that of the others.

Pasadena, 1988). This *Sgs-7/8* element has recently been shown to be capable of functioning as an enhancer when combined with the proximal element of *Sgs-3* (A. Hofmann, M. D. Garfinkel, and E. M. Meyerowitz, unpublished data). This indicates that the *Sgs-7/8* element is also functionally equivalent to the distal and proximal elements of *Sgs-3* and might also have the same regulatory sequence.

The sequence surrounding the distal element of *Sgs-3* and that surrounding the *Sgs-7/8* element were aligned with the sequence of the proximal element surrounding the two important regions which we have identified in this work and were searched for similarity. Figure 5 shows that a similar sequence is found in each of the three elements. The common sequence has been boxed. Note first of all that the identified homology extends beyond the 6 bp identified in the point mutational analysis but falls within the deletions which seriously affect expression. Second, there are two copies of the homology in the regulatory region of *Sgs-7/8*. Both have the same orientation, that is, facing *Sgs-7*. Third, the spacing between the two required regions of the homology is variable, ranging from 18 bp for the proximal element of *Sgs-3* to 39 bp for element 2 of *Sgs-7/8*. Each of these spacer regions contains a stretch of A+T-rich sequence.

The regulatory element consensus sequence. If we believe that the sequence homology identified within the regulatory elements of *Sgs-3*, *Sgs-7*, and *Sgs-8* is involved in tissue-specific expression, we would expect to find copies of it near other salivary gland-specific genes. Table 4 shows that this prediction is also confirmed. There are several copies of both the T-TTTG and the TCCA parts of the homology within the

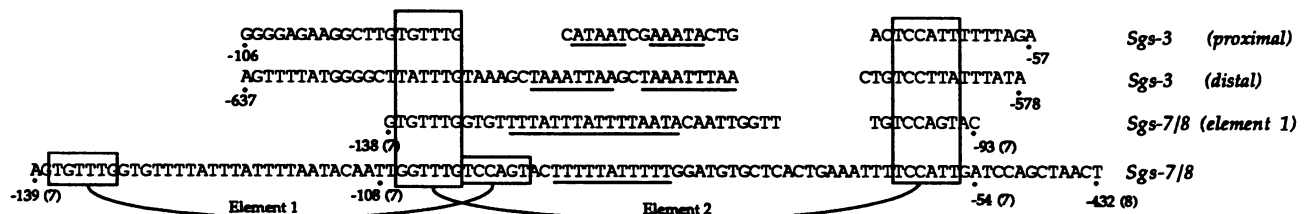


FIG. 5. Comparison of the sequence of the proximal element of *Sgs-3* with two other functionally related elements. Sequences of the proximal element of *Sgs-3*, the distal element of *Sgs-3*, and the regulatory region of *Sgs-7* and *Sgs-8* were compared. Blank spaces between sequences were inserted where needed to allow alignment. The regulatory region of *Sgs-7/8* was found to have two homologous elements. Each of the four elements of these three genes consists of two regions of homology, with a spacer region between. Underlines indicate A-T sequences in the spacer regions. Numbering, as for Fig. 1, is from the transcription start site of *Sgs-3*, *Sgs-7* (7), or *Sgs-8* (8).

upstream region of *Sgs-4*. Three pairs of these homologies have a spacing within the range of those seen at 68C and coincide with the three elements previously identified (17) as being involved in tissue-specific expression of this gene. Note that one of the listed *Sgs-4* elements has an orientation opposite to that of all the others. It is reasonable to assume that such an element could be functional because gene constructs in which the distal element of *Sgs-3* is in an inverted orientation still function, though at a reduced level (11, 43). The nucleotide sequence of the *Sgs-5* gene has also been determined (46), and we found the *Sgs-3* motif in the upstream region of this gene as well. Although these workers found that a deletion which removes the T-TTTG half of the motif still expresses, we would like to note that the sequence of the vector placed next to the breakpoint in the construct tested contains homology to the part of the motif which was deleted (pBR322: 5'-TCATGTTTGACAGCTTATCATCG| fusion). We also note that one of only three single-base-pair alterations found in the 5' DNA of a nonexpressed allele of *Sgs-5* lies very near the TCCA half of the motif. Three of the four homologies in these two genes also contain A+T-rich sequences in the spacer. From comparison of these sequences we have derived a consensus sequence, shown on the last line of Table 4. All of the base pairs of the consensus are conserved in *Drosophila simulans*, *Drosophila erecta*, and *Drosophila yakuba*, three species whose *Sgs-3* homologs are correctly expressed when transformed into *D. melanogaster* (27). Similarly, when the upstream sequence of the *Sgs-3* homolog *lgs-1* of *D. virilis* (a distantly related species) is searched, this consensus element is also found there (51).

The consensus sequence consists of two separated regions and a spacer sequence between them. The sequence is 5'-TNTTTG-spacing-TCCAT_A-3'. Spacings of 39 and 31 (*Sgs-7/8* elements), 28 (the distal element of *Sgs-3*), 18 (the proximal element of *Sgs-3*), 31 and 29 (*Sgs-4* elements), and 39 (*Sgs-5* element) bp allow for the specific expression of each gene. Furthermore, internal deletion mutant dm(-83/-75), which deletes 9 bp between the TGTTTG and the TCCATT sequences, directs a high level of expression (Fig. 2A). Therefore, a spacing of 9 bp is also allowable. These remaining base pairs are not particularly A+T rich, so the significance of this quality of the spacer region is unclear. In contrast, a 20-bp deletion in the spacer of the distal element expresses less than 50% of the level of the undelated gene (41).

Nature of the consensus element. The consensus element consists of two separated regions of homology. In the first of these regions, the bases TNTTTG are closely conserved. This is part of an imperfect inverted repeat, found at -76 to -90 bp of the proximal element, which is closely related to a sequence proposed as the ecdysterone response element of *hsp23* and also found upstream of a number of other ecdysterone-regulated genes (31). However, deletion of half of the inverted repeat does not seriously affect expression, so the symmetry of the homology may not be relevant. It shows no homology to any published steroid receptor binding site (see reference 7) nor to the ecdysterone response element identified for *hsp27* (42). This conserved region does have homology to the octamer motif of vertebrate genes, which is bound by a family of octamer-binding proteins (16), including a lymphoid specific factor, Oct-2 (3), and a factor expressed ubiquitously, Oct-1 (50). This homology is especially striking in the proximal element of *Sgs-3* studied here, where the sequence 5'-GTTTGCAT-3' is a 7 of 8 match to a consensus site found upstream of all immunoglobulin light

chain genes and of the herpesvirus *Tk* gene (37, 38), among other genes (see reference 16 for a review). The significance of this homology cannot be assessed on the basis of the evidence at hand, although the identity of this element as the binding site of a ubiquitous enhancer-like protein would fit well with observations from in vivo footprinting studies (Mathers, Ph.D. thesis [see model below]). In contrast, the TCCAT_A region shows no obvious homology to other known regulatory elements.

The two regions of homology are separated by nonconserved sequences (spacers) of various lengths. Most, but not all, of these spacers in the elements identified contain A+T-rich stretches. All but 9 bp of the spacer in the proximal element of *Sgs-3* can be deleted without seriously affecting function. When 20 bp of the spacer of the *Sgs-3* distal element is deleted (removing the A+T-rich sequence and leaving a 12-bp spacer), expression is reduced two- to threefold (41). The variety of allowable distances between the two homologies and the nonconservation of spacer sequence probably indicate that the length and particular sequence between the homologies do not have any important role in the functioning of the element. However, there is another possible interpretation. The spacing of each element can be classified into one of four groups by its length. From shorter to longer, they are as follows: (i) 9 bp (9-bp internal deletion of the proximal element of *Sgs-3*); (ii) 18 bp (the proximal element of *Sgs-3*); (iii) 28, 29, and 31 bp (the distal element of *Sgs-3*, element 1 of *Sgs-7/8*, and the *Sgs-4* elements, respectively); and (iv) 39 bp (element 2 of *Sgs-7/8* and the *Sgs-5* element). The length of the spacer increases by approximately 10 or 11 bp for each group from (i) to (iv). The activity of some genes is known to be a periodic function of the amount of DNA between the binding sites of its regulatory proteins. The period observed here is close to 10.5 bp, that is, one turn of the B-form DNA helix. These periodic phenomena have been attributed to a requirement for a unique rotational alignment in forming essential contacts between the bound proteins (4, 6, 19, 39). The 10- or 11-bp periodic spacing in the glue gene elements may be explained in the same manner.

Two domains within the proximal element. A remaining question is whether these two separated regions of sequence are recognized by a single protein or by two proteins. We postulate that two different proteins bind to the two conserved regions of each element for the following reasons. First, the fact that the spacer region sequence is not conserved and, in the case of the *Sgs-3* proximal element, can mostly be deleted argues that no specific protein interaction is occurring there and that the two conserved regions are acting independently. In addition, the periodicity of the spacing, as described above, argues for a protein-protein interaction. Second, lesions in the two regions have different quantitative effects on expression. Only one point mutation in the TCCAT_A sequence reduces expression, and deletion reduces expression by just over half, whereas five point mutations in the TNTTTG sequence lead to reduced expression, and deletion here leads to a fivefold reduction. Also, neither deletion alone leads to complete loss of function of the proximal element. Third, in vivo footprinting studies (Mathers, Ph.D. thesis) demonstrate that the two functionally important regions of conserved sequence are bound by proteins in vivo, but with different patterns of binding. In this study, chromatin prepared from a small number of dissected salivary glands or other tissues was treated with dimethyl sulfate and then purified, and the DNA of the proximal element was amplified by polymerase chain reac-

tion. The pattern of protected G residues in salivary glands of mid- to late third-instar larvae was compared with that in glands of early third-instar larvae and with that in prepupae, stages at which the *Sgs-3* gene is not expressed, as well as with the patterns in other tissues of the third-instar larva, such as fat body and gut. While site 1 (the region of the TNTTTG sequence) is bound in all tissues and times during the third instar, site 2 (the region of the TCCAT_A sequence) is bound only when the glue gene is being expressed. Furthermore, the latter site is not bound in the third-instar larval salivary glands of *npr-1*³ mutants or in the ecdysoneless mutant *l(1)su(f)^{ts67g}* (neither of which expresses *Sgs-3*), while the former site is bound in these mutants. This implies that at least two proteins are binding, each specific for one of the conserved sites, and therefore that there are two regulatory domains within what we have called the proximal element.

Interactions between the elements. It has frequently been observed that DNA binding proteins can interact with each other to give enhanced levels of expression or inducibility and that these interactions can be homologous (multiples of the same protein) or heterologous (different transcriptional activators) (16, 26). For some factors, this feature is so striking that it has been called promiscuous cooperativity (40). For instance, multiple copies of a glucocorticoid response element dramatically increase steroid inducibility of an attached gene (15, 52). It seems to be a common property of steroid-responsive genes that heterologous as well as homologous cooperations can contribute to the expression patterns observed. We have previously shown that the proximal and distal elements cooperate to direct high-level expression of the ecdysone-regulated *Sgs-3* gene (43). The present work (together with the work of Mathers, discussed above) allows us to propose a model for these interactions (Fig. 6). Each of the elements, proximal and distal, includes binding sites for two different factors. The factor which binds to site 1 appears to be present ubiquitously, whereas the factor which binds at site 2 is specific for the third-instar salivary gland (Mathers, Ph.D. thesis). We propose that the site 2 protein must interact with the site 1 enhancer-type protein in a heterologous interaction to allow stage- and tissue-specific expression; site 1 or site 2 alone gives no expression. In addition, site 1 protein bound at the proximal element interacts homologously with other site 1 proteins bound farther upstream to give synergistic enhancement of the level of expression. Since there might also be heterologous interactions with other site 2 proteins bound farther upstream, one might expect a fairly complex array of levels of expression, depending on how many binding sites for each of the two factors is present in a given locus. This is in fact what has been seen by us and others (28, 43): in constructs where upstream binding sites are still present, deleting part or all of sites 1 or 2 leaves varying amounts of residual activity. In this context it is interesting to note the point mutant PM(-90). This single-base-pair change has the effect of completely inactivating the proximal element, equivalent to deleting both site 1 and site 2. One possible explanation for this could be that this mutant site is still bound by the site 1 protein, but the protein is now in a conformation which prevents proteins bound at the proximal element from interacting with proteins bound farther upstream.

This analysis has allowed us to identify at least the core of the binding sites for two different regulatory factors of the *Sgs-3* glue gene. With this information, we have started a search for the DNA binding proteins which act through these sites.

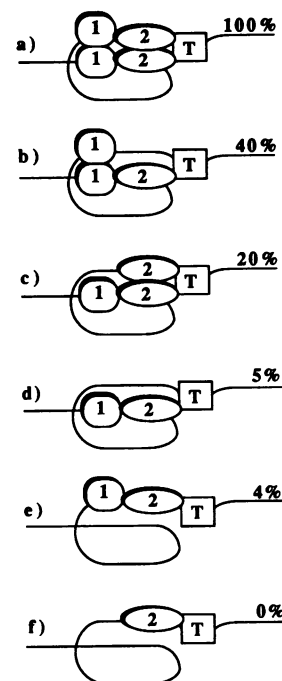


FIG. 6. A possible model for the interactions of site 1 and site 2 binding proteins with the proximal and distal regulatory elements of *Sgs-3*. Mutations in either binding site of the proximal element still allow expression, though at a lower level, because of interactions of the proteins bound to the deleted element with proteins bound to the distal element upstream. (a) through (e) Percentages represent levels of expression and are based on ADH enzyme assays of salivary gland extracts from animals transformed with various *Sgs-3-Adh* deletion constructs (this report and reference 43). f) The 0% is based on results with an *Sgs-7-Adh* deletion construct with 92 bp of 5' sequence, including a single site 2 homologous sequence, which does not express any detectable ADH enzyme activity (Hofmann et al., unpublished data). Symbols: 1 and 2, proteins which bind to sites 1 and 2, respectively, of the proximal regulatory element (see Discussion); T, TATA-binding factor (see reference 10).

ACKNOWLEDGMENTS

The order of the first three authors on the title page is arbitrary. We thank Horst Kress for generously sharing data with us prior to publication. We are grateful to Mark D. Garfinkel, Annemarie Hofmann, and Peter H. Mathers for helpful discussions during the course of this work and to Bruce A. Hamilton and Michael J. Palazzolo for critical reading of the manuscript.

This work was supported in part by a Public Health Service grant (GM28075) to E.M.M. from the Institute for General Medical Sciences, National Institutes of Health, and by a generous grant from the Lucille P. Markey Charitable Trust. T.T. was supported by a grant from the Procter and Gamble Co., and M.R. was supported by a National Research Service Award (GM 10694).

LITERATURE CITED

1. Akam, M. E., D. B. Roberts, G. P. Richards, and M. Ashburner. 1978. *Drosophila*: the genetics of two major larval proteins. *Cell* 13:215-225.
2. Bender, W., P. Spierer, and D. S. Hogness. 1983. Chromosomal walking and jumping to isolate DNA from the *Ace* and *rosy* loci and the *bithorax* complex in *Drosophila melanogaster*. *J. Mol. Biol.* 168:17-33.
3. Clerc, R. G., L. M. Corcoran, J. H. LeBowitz, D. Baltimore, and P. A. Sharp. 1988. The B-cell-specific Oct-2 protein contains POU box- and homeo box-type domains. *Genes Dev.* 2:1570-1581.
4. Cohen, R. S., and M. Meselson. 1988. Periodic interaction of

- heat shock transcriptional elements. *Nature* (London) **332**:856–857.
5. Crowley, T. E., P. H. Mathers, and E. M. Meyerowitz. 1984. A *trans*-acting regulatory product necessary for expression of the *Drosophila melanogaster* 68C glue gene cluster. *Cell* **39**:149–156.
 6. Dunn, T. M., S. Hahn, S. Ogden, and R. F. Schleif. 1984. An operator at –280 base pairs that is required for expression of araBAD operon promoter: addition of DNA helical turns between the operator and promoter cyclically hinders repression. *Proc. Natl. Acad. Sci. USA* **81**:5017–5020.
 7. Evans, R. M. 1988. The steroid and thyroid hormone receptor superfamily. *Science* **240**:889–895.
 8. Fraenkel, G., and V. J. Brookes. 1953. The process by which the puparia of many species of flies become fixed to a substrate. *Biol. Bull.* **105**:442–449.
 9. Garfinkel, M. D., R. E. Pruitt, and E. M. Meyerowitz. 1983. DNA sequences, gene regulation and modular protein evolution in the *Drosophila* 68C glue gene cluster. *J. Mol. Biol.* **168**:765–789.
 10. Giangrande, A., C. Mettling, M. Martin, C. Ruiz, and G. Richards. 1989. *Drosophila Sgs3* TATA: effects of point mutations on expression *in vivo* and protein binding *in vitro* with staged nuclear extracts. *EMBO J.* **8**:3459–3466.
 11. Giangrande, A., C. Mettling, and G. Richards. 1987. *Sgs-3* transcript levels are determined by multiple remote sequence elements. *EMBO J.* **6**:3079–3084.
 12. Goodbourn, S., and T. Maniatis. 1988. Overlapping positive and negative regulatory domains of the human β -interferon gene. *Proc. Natl. Acad. Sci. USA* **85**:1447–1451.
 13. Guild, G. 1984. Molecular analysis of a developmentally regulated gene which is expressed in the larval salivary gland of *Drosophila*. *Dev. Biol.* **102**:462–470.
 14. Hansson, L., and A. Lambertsson. 1983. The role of *su(f)* gene function and ecdysterone in transcription of glue polypeptide mRNAs in *Drosophila melanogaster*. *Mol. Gen. Genet.* **192**:395–401.
 15. Jantzen, H.-M., U. Strähle, B. Gloss, F. Stewart, W. Schmid, M. Boshart, R. Miksicek, and G. Schütz. 1987. Cooperativity of glucocorticoid response elements located far upstream of the tyrosine aminotransferase gene. *Cell* **49**:29–38.
 16. Johnson, P. F., and S. L. McKnight. 1989. Eukaryotic transcriptional regulatory proteins. *Annu. Rev. Biochem.* **58**:799–839.
 17. Jongens, T. A., T. Fowler, A. W. Sheromoen, and S. K. Beckendorf. 1988. Functional redundancy in the tissue-specific enhancer of the *Drosophila Sgs-4* gene. *EMBO J.* **7**:2559–2567.
 18. Korge, G. 1975. Chromosome puff activity and protein synthesis in larval salivary glands of *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* **72**:4550–4554.
 19. Kramer, H., M. Niemoller, M. Amouyal, B. Revert, B. von Wilcken-Bergmann, and B. Muller-Hill. 1987. *lac* repressor forms loops with linear DNA carrying two suitably spaced *lac* operators. *EMBO J.* **6**:1481–1491.
 20. Kunkel, T. A. 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl. Acad. Sci. USA* **82**:488–492.
 21. Kunkel, T. A., J. D. Roberts, and R. A. Zakour. 1987. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Methods Enzymol.* **154**:367–382.
 22. Landschultz, W. H., P. F. Johnson, E. Y. Adashi, B. J. Graves, and S. L. McKnight. 1988. Isolation of a recombinant copy of the gene encoding C/EBP. *Genes Dev.* **2**:786–800.
 23. Levy, D. E., D. S. Kessler, R. Pine, N. Reich, and J. E. Darnell, Jr. 1988. Interferon-induced nuclear factors that bind a shared promoter element correlate with positive and negative transcriptional control. *Genes Dev.* **2**:383–393.
 24. Lowry, O. D., N. J. Rosebrough, A. L. Farr, and R. J. Randall. 1951. Protein measurements with the folin phenol reagent. *J. Biol. Chem.* **193**:265–275.
 25. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
 26. Maniatis, T., S. Goodbourn, and J. A. Fischer. 1987. Regulation of inducible and tissue-specific gene expression. *Science* **236**:1237–1245.
 27. Martin, C. M., C. A. Mayeda, and E. M. Meyerowitz. 1988. Evolution and expression of the *Sgs-3* glue gene of *Drosophila*. *J. Mol. Biol.* **201**:273–287.
 28. Martin, M., A. Giangrande, C. Ruiz, and G. Richards. 1989. Induction and repression of the *Drosophila Sgs-3* glue gene are mediated by distinct sequences in the proximal promoter. *EMBO J.* **8**:561–568.
 29. Martin, M., C. Mettling, A. Giangrande, C. Ruiz, and G. Richards. 1989. Regulatory elements and interactions in the *Drosophila* 68C glue gene cluster. *Dev. Genet.* **10**:189–197.
 30. Martin, P., A. Martin, A. Osmani, and W. Sofer. 1986. A transient expression assay for tissue-specific gene expression of alcohol dehydrogenase in *Drosophila*. *Dev. Biol.* **117**:574–580.
 31. Mestrlil, R., P. Schiller, J. Amin, H. Klapper, J. Ananthan, and R. Voellmy. 1986. Heat shock and ecdysterone activation of the *Drosophila melanogaster hsp23* gene: a sequence element implied in developmental regulation. *EMBO J.* **5**:1667–1673.
 32. Meyerowitz, E. M., and D. S. Hogness. 1982. Molecular organization of a *Drosophila* puff site that responds to ecdysone. *Cell* **28**:165–176.
 33. Meyerowitz, E. M., K. Vijay Raghavan, P. H. Mathers, and M. Roark. 1987. How *Drosophila* larvae make glue: control of *Sgs-3* gene expression. *Trends Genet.* **3**:288–293.
 34. Muskavitch, M., and D. S. Hogness. 1980. Molecular analysis of a gene in a developmentally regulated puff of *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* **77**:7362–7366.
 35. Myers, R. M., L. S. Lerman, and T. Maniatis. 1985. A general method for saturation mutagenesis of cloned DNA fragments. *Science* **229**:242–247.
 36. Myers, R. M., K. Tilly, and T. Maniatis. 1986. Fine structure genetic analysis of a β -globin promoter. *Science* **232**:613–618.
 37. Parslow, T. G., D. L. Blair, W. J. Murphy, and D. K. Granner. 1984. Structure of the 5' ends of immunoglobulin genes: a novel conserved sequence. *Proc. Natl. Acad. Sci. USA* **81**:2650–2654.
 38. Parslow, T. G., S. D. Jones, B. Bond, and K. R. Yamamoto. 1987. The immunoglobulin octanucleotide: independent activity and selective interaction with enhancers. *Science* **235**:1498–1501.
 39. Ptashne, M. 1986. Gene regulation by proteins acting nearby and at a distance. *Nature* (London) **322**:697–701.
 40. Ptashne, M. 1988. How eukaryotic transcriptional activators work. *Nature* (London) **335**:683–689.
 41. Romain, P., A. Giangrande, G. Richards, and M. Bellard. 1988. Analysis of a DNase I-hypersensitive site in transgenic *Drosophila* reveals a key regulatory element of *Sgs-3*. *Proc. Natl. Acad. Sci. USA* **85**:2718–2722.
 42. Riddihough, G., and H. R. B. Pelham. 1987. An ecdysone element in the *Drosophila hsp27* promoter. *EMBO J.* **6**:3729–3734.
 43. Roark, M., K. Vijay Raghavan, T. Todo, C. Mayeda, and E. M. Meyerowitz. 1990. Cooperative enhancement at the *Drosophila Sgs-3* locus. *Dev. Biol.* **139**:121–133.
 44. Rubin, G. M., and A. C. Spradling. 1982. Genetic transformation of *Drosophila* with transposable element vectors. *Science* **218**:348–353.
 45. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463–5467.
 46. Shore, E. M., and G. M. Guild. 1987. Closely linked DNA elements control the expression of the *Sgs-5* glue protein gene in *Drosophila*. *Genes Dev.* **1**:829–839.
 47. Singh, H., R. G. Clerc, and J. H. Lebowitz. 1989. Molecular cloning of sequence-specific DNA binding proteins using recognition site probes. *Biotechniques* **7**:252–261.
 48. Sofer, W., and H. Ursprung. 1968. *Drosophila* alcohol dehydrogenase: purification and partial characterization. *J. Biol. Chem.* **243**:3110–3115.
 49. Spradling, A. C., and G. M. Rubin. 1982. Transposition of cloned P elements into *Drosophila* germ line chromosomes. *Science* **218**:341–347.
 50. Sturm, R. A., G. Das, and W. Herr. 1988. The ubiquitous

- octamer-binding protein Oct-1 contains a POU domain with a homeo box subdomain. *Genes Dev.* **2**:1582–1599.
51. **Swida, U., L. Lucka, and H. Kress.** 1990. Glue protein genes in *Drosophila virilis*: their organization, developmental control of transcription, and specific mRNA degradation. *Development* **108**:269–280.
52. **Tsai, S. Y., M.-J. Tsai, and B. O. O'Malley.** 1989. Cooperative binding of steroid hormone receptors contributes to transcriptional synergism at target enhancer elements. *Cell* **57**:443–448.
53. **Vieira, J., and J. Messing.** 1987. Production of single-strand plasmid DNA. *Methods Enzymol.* **153**:3–11.
54. **Vijay Raghavan, K., M. A. Crosby, P. H. Mathers, and E. M. Meyerowitz.** 1986. Sequences sufficient for correct regulation of Sgs-3 lie close to or within the gene. *EMBO J.* **5**:3321–3326.
55. **Weiher, H., M. König, and P. Gruss.** 1983. Multiple point mutations affecting the Simian Virus 40 enhancer. *Science* **219**:626–631.
56. **Wood, W. I., J. Gitsheier, L. A. Lasky, and R. M. Lawn.** 1985. Base composition-independent hybridization in tetramethylammonium chloride: a method for oligonucleotide screening of highly complex gene libraries. *Proc. Natl. Acad. Sci. USA* **82**:1585–1588.